2

RADC-TR-90-226
In-House Report
November 1990



# MULTICHANNEL LINEAR PREDICTION AND ITS ASSOCIATION WITH TRIANGULAR MATRIX DECOMPOSITION



James H. Michels



**Rome Air Development Center**
**Air Force Systems Command**
**Griffiss Air Force Base, NY 13441-5700**

This report has been reviewed by the RADC Public Affairs Division (PA) and is releasable to the National Technical Information Services (NTIS) At NTIS it will be releasable to the general public, including foreign nations.

RADC-TR-90-226 has been reviewed and is approved for publication.

APPROVED:

JOSEPH J. POLNIASZEK
Acting Chief, Surveillance Technology Division
Directorate of Surveillance

APPROVED:

JAMES W. YOUNGBERG, Lt Col, USAF
Deputy Director of Surveillance

FOR THE COMMANDER:

JAMES W. HYDE III
Directorate of Plans & Programs

If your address has changed or if you wish to be removed from the RADC mailing list, or if the addressee is no longer employed by your organization, please notify RADC ( OCTM) Griffiss AFB NY 13441-5700. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

# REPORT DOCUMENTATION PAGE

Form Approved
OPM No. 0704-0188



1. AGENCY USE ONLY *(Leave Blank)*

2. REPORT DATE
November 1990

3. REPORT TYPE AND DATES COVERED
In-House Apr 89 - May 90

4. TITLE AND SUBTITLE
MULTICHANNEL LINEAR PREDICTION AND ITS ASSOCIATION WITH TRIANGULAR MATRIX DECOMPOSITION

5. FUNDING NUMBERS
PE - 62702F
PR - 4506
TA - 17
WU - 67

6. AUTHOR(S)
James H. Michels

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Rome Air Development Center (OCTM)
Griffiss AFB NY 13441-5700

8. PERFORMING ORGANIZATION REPORT NUMBER
RADC-TR-90-226

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Rome Air Development Center (OCTM)
Griffiss AFB NY 13441-5700

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
RADC Project Engineer: James H. Michels/OCTM/(315) 330-4431

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The determination of the multichannel coefficients of a minimum mean-square error prediction filter from covariance information is considered in this report. The Levinson recursion for computing these coefficients is examined in the 1-D multichannel case. Specifically, the multichannel coefficients of linear prediction for random processes and the prediction error covariance matrix are related to the multichannel covariance matrix of the observation data by triangular decomposition.

14. SUBJECT TERMS
Linear Prediction
Prediction Error Filters
Multichannel Random Processes
Adaptive Filtering
Parameter Estimation

15. NUMBER OF PAGES
40

16. PRICE CODE

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| UNCLASSIFIED | UNCLASSIFIED | UNCLASSIFIED | UL |

NSN 7540-01-280-5500

Standard Form 298 880922
Prescribed by ANSI Std. Z39-18
298-01

## ACKNOWLEDGEMENTS

The author wishes to express his appreciation to Dr. K. S. Arun for his interest and review of this work and, in addition, for indicating the similarity of multichannel one-dimensional processing to that of the single channel, two dimensional case [9].

## CONTENTS

I. Introduction

In this report, multichannel, one-dimensional linear prediction is considered where the zero-mean observation processes are assumed to be generated by a multivariate autoregressive model with a white noise driving vector $\underline{u}(n)$ that is partially correlated over the channels. It is proposed here that this white noise driving vector can be modeled in terms of a temporally and 'spatially' (ie., over the channels) causal vector $\underline{z}(n)$ uncorrelated across channels such that $\underline{u}(n)=L_u \underline{z}(n)$ where $L_u$ is unit diagonal lower triangular. A one-dimensional multichannel prediction error filter of order p is initially considered with coefficients of the form $A_p^H(k)$, k=0,1,...,p where $A_p^H(k)$ is the kth matrix coefficient obtained from the multichannel normal equation. The output error vector $\underline{\varepsilon}_p(n)$ of this filter has a non-diagonal forward error covariance matrix ${}^pR_{\varepsilon\varepsilon}(0)\equiv[\Sigma_f]_p$ which is decomposed by an $L_{\gamma_p} D_{\gamma_p} L_{\gamma_p}^H$ decomposition. It is shown that $L_{\gamma_p}^{-1}$ is an estimate of $L_u^{-1}$ and that a filter with coefficients $L_{\gamma_p}^{-1} A_p^H(k)$, k=0,1,...,p yields an output $\underline{\gamma}_p(n)$ which equals $\underline{z}(n)$ when the coefficient estimates are minimum mean-square error (MMSE). Thus, a two stage filtering operation is implemented first to obtain the coefficients $A_p^H(k)$, k=0,1,...,p and $[\Sigma_f]_p$ followed by the $L_{\gamma_p} D_{\gamma_p} L_{\gamma_p}^H$ decomposition of $[\Sigma_f]_p$ to obtain $L_{\gamma_p}^{-1}$. Finally, it is shown that the coefficients of the form $L_{\gamma_p}^{-1} A_p^H(k)$, k=0,1,...,p and p=0,1,...,P are contained in the block rows of $L_{\underline{\beta}}^{-1}$ where $L_{\underline{\beta}} D_{\underline{\beta}} L_{\underline{\beta}}^H$ is a decomposition of the multichannel correlation matrix $R_{\underline{xx}}$ written in index ordered form, $L_{\underline{\beta}}$ is unit diagonal lower triangular and $D_{\underline{\beta}}$ is strictly diagonal (ie., with 1x1 diagonal elements). Thus, the pth block row of $L_{\underline{\beta}}^{-1}$ contains the coefficients of the pth order multichannel prediction error filter. This result implies that a set of

coefficients equivalent to those of $L\gamma_p^{-1} A_p^H(k)$, k=0,1,...,p could be obtained in a single stage of computation.

In [1,14], a model-based multichannel likelihood ratio is presented which considers the detection of a signal vector in an additive non-white noise vector disturbance process. The implementation of this detection scheme utilizes prediction error filters to whiten the observation processes in time and across channels and provides a practical framework to utilize the coefficient estimates. In addition, the whitening operation in time and across channels provides a generalization of the innovations-based likelihood ratio suggested in [11,12] to the mutichannel case. As noted in [13], this approach provides a generalization of several detection problems as well as a practical implementation architecture.

It is anticipated that recognition of the result presented here will motivate the development of computationally efficient methods to obtain the multichannel prediction error coefficients. Furthermore, the similarity of the processing methods discussed here to those used in two dimensional image processing is noted in [9].

## II. Autoregressive Process Models

In this analysis, the observation processes are assumed to be multichannel autoregressive. The order M autoregressive (AR) model for a Jx1 vector process $\underline{x}(n)$ is

$$\underline{x}(n) = -\sum_{k=1}^{M} A_M^H(k)\underline{x}(n-k) + \underline{u}(n) \tag{2.1}$$

where $A_M^H(k)$ is the kth JxJ matrix coefficient for the AR process and H denotes the Hermitian operation (ie., the complex conjugate transpose operation). We note that $A_M^H(k)$ is expressed in terms of the Hermitian operation for notational convenience, but is not necessarily a Hermitian matrix. The vector $\underline{u}(n)$ is a Jx1 white noise driving vector which, in general, has an arbitrary correlation across the J channels so that

$$E[\underline{u}(n)\underline{u}^H(n-l)] = \begin{cases} [0] & l \neq 0 \\ R_{uu}(0) \equiv [\Sigma_f]_M & l = 0. \end{cases} \tag{2.2}$$

where $R_{uu}(0) \equiv [\Sigma_f]_M$ is the JxJ positive semi-definite covariance matrix of the forward white noise driving vector $\underline{u}(n)$ and may have off-diagonal components. Since $\underline{u}(n)$ is uncorrelated in time, but retains an arbitrary correlation across channels, then with wide-sense joint stationarity of the channel processes assumed, we can consider

$$\underline{u}(n) = C_v\underline{v}(n) \tag{2.3}$$

where the J x J matrix $C_v$ is a constant matrix, and $\underline{v}(n)$ is a white noise vector uncorrelated both in time and across channels such that

$$E[\underline{v}(n)\underline{v}^H(n-l)] = \begin{cases} [0] & l \neq 0 \\ D_v & l = 0. \end{cases} \tag{2.4}$$

The elements of the diagonal matrix $D_V$ are the variance terms associated with the white noise driving term on each channel. And so, from eq (2.3) we can obtain the zero-lag correlation matrix

$$R_{uu}(0) = E\,[\underline{u}(n)\underline{u}^H(n)] \tag{2.5a}$$

$$= E\,[C_v\underline{v}(n)\,\underline{v}^H(n)C_v^H] \tag{2.5b}$$

$$= C_v D_V C_v^H. \tag{2.5c}$$

We could assume unit variance on all elements of $D_V$ without loss of generality; ie., $D_V=I$, the JxJ identity matrix, so that eq(2.5c) is the Cholesky decomposition $R_{uu}(0)=C_vC_v^H$ when $C_v$ is lower triangular. Since $R_{uu}(0)$ expressed in eq(2.5) is Hermitian† , positive semi-definite, we could also consider an $LDL^H$ decomposition* of $R_{uu}(0)$ such that

$$R_{uu}(0) = L_u D_u L_u^H \tag{2.6}$$

where $L_u$ is unit diagonal lower triangular and $D_u$ is strictly diagonal with non-negative entries only. Assuming $D_V=I$, the matrices $L_u$ and $C_v$ are related as

$$C_vC_v^H = L_u D_u L_u^H \tag{2.7}$$

so that

$$C_v = L_u D_u^{1/2}. \tag{2.8}$$

Solving eq(2.6) for $D_u$, we obtain

---

† It is noted that in general the correlation matrix $R_{uu}(l)$ is not Hermitian for $l \neq 0$.

* The motivation for considering the $LDL^H$ decomposition stems from its relation to linear prediction [3] and the desire to determine both temporal and cross-channel whitening with a single stage of decomposition.

$$D_u = L_u^{-1} R_{uu}(0)(L_u^{-1})^H \tag{2.9a}$$

$$= E\,[L_u^{-1}\underline{u}(n)\underline{u}^H(n)(L_u^{-1})^H] \tag{2.9b}$$

$$= E\,[\underline{z}(n)\underline{z}^H(n)] \tag{2.9c}$$

where

$$\underline{z}(n) = L_u^{-1}\underline{u}(n) \tag{2.10}$$

is a J x 1 vector containing uncorrelated elements. Note that from eqs(2.3), (2.8) and (2.10), we can obtain

$$\underline{z}(n) = L_u^{-1} C_v\, \underline{v}(n) \tag{2.11a}$$

$$= D_u^{1/2}\underline{v}(n) \tag{2.11b}$$

when $D_v$=I. Thus $\underline{z}(n)$ represents a temporally and "spatially-causal" white noise driving term to the multichannel AR model. Since $L_u^{-1}$ is also lower triangular unit diagonal, it is invertible so that from eq(2.10)

$$\underline{u}(n) = L_u\underline{z}(n). \tag{2.12}$$

Eq (2.12) indicates that $\underline{u}(n)$, originally defined in eq(2.3), could identically be generated by the $\underline{z}(n)$ process through the transformation matrix $L_u$; i.e., eq(2.1) can be written in the equivalent form

$$\underline{x}(n) = -\sum_{k=1}^{M} A_M^H(k)\underline{x}(n-k) + L_u\,\underline{z}(n) \tag{2.13}$$

In chapter III, a two stage multichannel prediction error filter is considered which uses estimates of the $A_M^H(k)$ coefficients to obtain $\underline{u}(n)$ in the first stage of filtering and an estimate of $L_u^{-1}$ to obtain the temporally and spatially uncorrelated process $\underline{z}(n)$ in the second stage.

Finally, we note that the relationship between the matrix coefficients $A_M^H(k)$, the covariance matrix $[\Sigma_f]_M$ of the forward AR driving noise vector and the correlation matrix $R_{\underline{xx}}$ is expressed by the multichannel Yule-Walker equation; ie.,

$$\underline{A}_M^H[\tilde{R}_{\underline{xx}}] = \{[\Sigma_f]_M^H\ [0]\ ...[0]\} \tag{2.14}$$

where

$$\underline{A}_M^H = [I\ A_M^H(1)\ A_M^H(2)\ ...\ A_M^H(M)]. \tag{2.15a}$$

$$R_{\underline{xx}} = E[\underline{\underline{x}}_{n,n-M}\ \underline{\underline{x}}_{n,n-M}^H] \tag{2.15b}$$

$$\underline{\underline{x}}_{n,n-M}^T = [\underline{x}^T(n)\ \underline{x}^T(n-1)...\underline{x}^T(n-M)] \tag{2.15c}$$

and

$$\underline{x}^T(k) = [x_1(k)\ x_2(k)...x_J(k)]. \tag{2.15d}$$

The matrix $[\tilde{R}_{\underline{xx}}]$ is the reversed order correlation matrix of $[R_{\underline{xx}}]$; i.e., the correlation matrix obtained with the time order of the vector $\underline{\underline{x}}_{n,n-M}$ from eq(2.15c) reversed [10].

III. Multichannel Prediction Error Filtering

The output of a multichannel linear prediction error filter of order p is expressed as

$$\underline{\varepsilon}_p(n) = \underline{x}(n) - \hat{\underline{x}}(n|n-1) \tag{3.1a}$$

$$= \underline{x}(n) + \sum_{k=1}^{p} A_p^H(k)\, \underline{x}(n-k) \tag{3.1b}$$

$$= \sum_{k=0}^{p} A_p^H(k)\, \underline{x}(n-k) \tag{3.1c}$$

where $A_p^H(k)$ is the kth matrix coefficient of the linear predictor with k=1,2,...,p and $A_p^H(0)=I$, the JxJ identity matrix. The subscript p distinguishes the matrix coefficients and the error output as those from a filter of order p. In Appendix A, it is shown that when the matrix coefficients in eq(3.1) satisfy the multichannel normal equations,

$$E\,[\underline{\varepsilon}_p(n)\, \underline{\varepsilon}_p{}^H(n-k)] = [0] \qquad k>0, \tag{3.2}$$

where the output vector process $\underline{\varepsilon}_p(n)$ is the MMSE prediction error. The multichannel normal equations for a prediction error filter of order p are expressed as

$$\underline{A}_p^H\, [\tilde{R}_{\underline{xx}}] = \{[\Sigma_f]_p^H\ [0]\ \ldots\ [0]\} \tag{3.3a}$$

where

$$\underline{A}_p^H = [I\ \ A_p^H(1)\ A_p^H(2)\ \ \ A_p^H(p)] \tag{3.3b}$$

$$[\tilde{R}_{\underline{xx}}] = E[\underline{\underline{X}}_{n,n-p}\ \underline{\underline{X}}_{n,n-p}] = E[\underline{\underline{x}}_{n-p,n}\ \underline{\underline{x}}^H_{n-p,n}] \tag{3.3c}$$

and

$$[\Sigma_f]_p = E[\underline{\varepsilon}_p(n)\underline{\varepsilon}_p{}^H(n)] = {}^pR_{\varepsilon\varepsilon}(0). \tag{3.3d}$$

$[\underline{\tilde{R}_{xx}}]$ is the reversed order correlation matrix defined in the previous section. We also note that eqs (3.3a) and (2.14) are identical in form. This equality implies that the MMSE estimate of the AR observation process is obtained when the prediction error filter coefficients are identically equal to the AR process coefficients. Assuming the applicability of the AR process as a representative model of the observation process, then as the coefficients $A_p^H(k)$ k=1,2,...,p approach the AR coefficients of the model; ie., for p≥M and

$$A_p^H(k) = \begin{cases} A_M^H(k) & k \le M \\ [0] & k > M \end{cases} \tag{3.4}$$

we have from eqs(2.1) and (3.1a)

$$\underline{\varepsilon}_p(n) = \underline{u}\,(n). \tag{3.5}$$

Figure 3.1 shows the synthesis and analysis procedure. Under the condition that the prediction error coefficients are <u>identical</u> to the coefficients of the AR model process and under the <u>assumption</u> that the AR process is the <u>exact</u> model of the observed process, the prediction error filter output $\underline{\varepsilon}_p(n)$ is a white noise vector equivalent to the AR model white noise driving vector. However, it must be emphasized that the use of an AR process with a white noise driving function is usually an approximate representation; i.e. it is not used to describe the underlying physical mechanisms which give rise to the random processes. Rather, it is an approximate model representation for these processes. We must therefore make a <u>distinction</u> between the model of the processes (synthesis) and the estimation process (analysis) [10]. In general, the output $\underline{\varepsilon}_p(n)$ of the linear predictor is not a white noise vector output due to the approximate representation of physical processes by an AR model as well as the fact that we often do not have 'a priori' knowledge regarding the values of the coefficients of this approximate model. As a result, we must estimate the filter coefficients from the observation data as we obtain it. With a limited amount of data, the filter coefficients are

only estimates of the AR process coefficients. Furthermore, assumptions such as the ergodic approximation are often made in the estimation of process parameters. In [2], the validity of this approximation is considered in terms of basic parameters of the observation processes. These results provide a clearer understanding as well as a means of determining the processing requirements needed to achieve parameter estimates with high confidence limits.

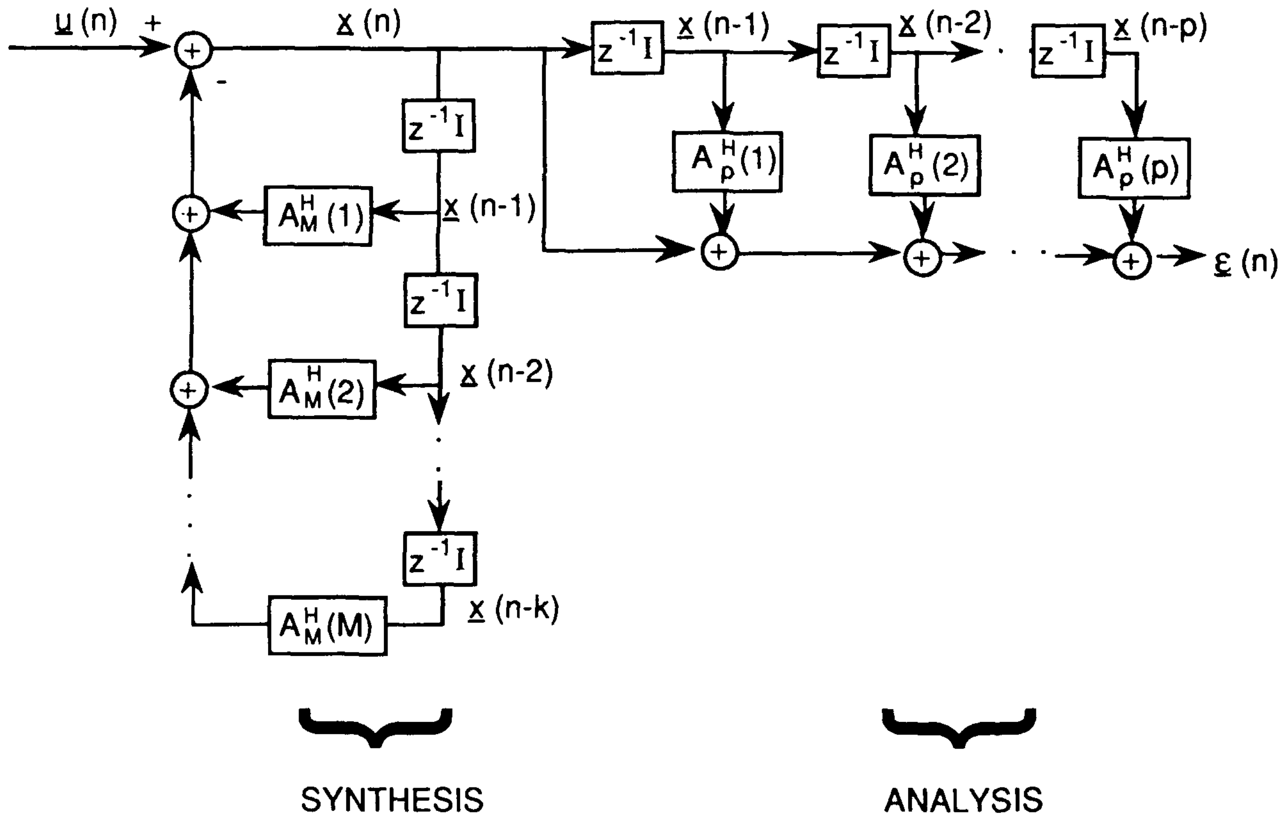


Figure 3.1

For stationary processes, these coefficients could be determined through estimates of the multichannel correlation matrix and the Levinson-Wiggins-Robinson algorithm [4]. Other methods proposed by Strand-Nuttall and Morf-Vieira have been developed [5,6,7,8] with improved performance with limited data. For non-stationary processes, adaptive schemes must be considered. We will address this topic in a subsequent report.

At this point, we note that eq(3.5) resulted from the analysis procedure [via a linear prediction error filter with coefficients given by eq(3.4)] of the process synthesized by eq (2.1). If $\underline{u}(n)$ is assumed to be uncorrelated across channels, the resulting $\underline{\varepsilon}_p(n)$ is also uncorrelated in time and space (i.e., channels). In general, however, $\underline{u}(n)$ may possess arbitrary correlation between the J channel elements. Therefore, the vector $\underline{\varepsilon}_p(n)$ will retain a residual correlation over the channels due to the spatial correlation of $\underline{u}(n)$.

Since the matrix ${}^pR_{\varepsilon\varepsilon}(0) = E\,[\underline{\varepsilon}_p(n)\underline{\varepsilon}_p^H(n)]$ is Hermitian†, and positive semi-definite, we can perform an $LDL^H$ decomposition* of ${}^pR_{\varepsilon\varepsilon}(0)$ such that

$$ {}^pR_{\varepsilon\varepsilon}(0) = L_{\gamma_p} D_{\gamma_p} L_{\gamma_p}^H . \tag{3.6}$$

Solving for $D_\gamma$

$$ D_{\gamma_p} = L_{\gamma_p}^{-1} R_{\varepsilon\varepsilon}(0)(L_{\gamma_p}^{-1})^H \tag{3.7a}$$

$$ = L_{\gamma_p}^{-1} E\,[\underline{\varepsilon}_p(n)\, \underline{\varepsilon}_p^H(n)]\, (L_{\gamma_p}^{-1})^H \tag{3.7b}$$

$$ = E\,[L_{\gamma_p}^{-1}\, \underline{\varepsilon}_p(n)\, \underline{\varepsilon}_p^H(n)\, (L_{\gamma_p}^{-1})^H] \tag{3.7c}$$

$$ = E\,[\underline{\gamma}_p(n)\, \underline{\gamma}_p^H(n)] \tag{3.7d}$$

where

$$ \underline{\gamma}_p(n) = L_{\gamma_p}^{-1}\, \underline{\varepsilon}_p(n) \tag{3.8}$$

so that the vector $\underline{\gamma}_p(n)$ contains uncorrelated elements. Also,

---

† It is noted that in general the correlation matrix $R_{\varepsilon\varepsilon}(l)$ is not Hermitian for $l \neq 0$.

* Other decompositions could be used such as Cholesky or unitary forms. However, as noted previously, the motivation for using the $LDL^H$ decomposition is to note its relation to linear prediction.

$$E\,[\underline{\varepsilon}_p(n)\underline{\varepsilon}_p^H(n')] = [0] \qquad n \neq n' \tag{3.9}$$

from the orthogonality principle. Then, using eq(3.8) to solve for $\underline{\varepsilon}_p(n)$ and substituting this result in (3.9), we obtain

$$E\,[L_{\gamma_p}\, \underline{\gamma}_p(n)\, \underline{\gamma}_p^H(n')\, L_{\gamma_p}^H] = [0] \qquad n \neq n' \tag{3.10}$$

so that

$$L_{\gamma_p} E\,[\, \underline{\gamma}_p(n)\, \underline{\gamma}_p^H(n')\,]\, L_{\gamma_p}^H = [0] \qquad n \neq n'. \tag{3.11}$$

Finally, since $L_{\gamma_p}$ is non-singular,

$$E\,[\underline{\gamma}_p(n)\, \underline{\gamma}_p^H(n')\,] = [0] \qquad n \neq n'. \tag{3.12}$$

Eq (3.12) implies that $\underline{\gamma}_p(n)$ retains its temporal decorrelation while eq(3.7d) denotes its spatial decorrelation. When

$$A_p^H(k) \approx \begin{cases} A_M^H(k) & k \leq M \\ [0] & k > M \end{cases} \tag{3.13}$$

the output of the first filter stage converges toward $\underline{u}(n)$, so that eq (3.8) becomes (noting that $R_{\varepsilon\varepsilon}(0) \approx R_{uu}(0)$ and the uniqueness of the $LDL^H$ decomposition)

$$\underline{\gamma}_p(n) = L_{\gamma_p}^{-1}\, \underline{\varepsilon}_p(n) \approx L_u^{-1}\, \underline{u}(n) \tag{3.14a}$$

$$\approx L_u^{-1} L_u \underline{z}(n) \tag{3.14b}$$

$$\approx \underline{z}(n) \tag{3.14c}$$

where eq(2.12) was used to obtain eq(3.14b). Thus, as the filter coefficients converge to the coefficients of the AR process, $\underline{\gamma}_p(n)$ approximates the spatially

and temporally whitened process $\underline{z}(n)$. In addition, $\underline{\gamma}_p(n)$ has been obtained through a causal and causally invertible transformation of the original observation process $\underline{x}(n)$; ie., the input vector $\underline{x}(n)$ could be recovered through an inverse filter operation on $\underline{\gamma}_p(n)$.

The estimate $\hat{L}_u^{-1}$ could be obtained by first estimating the correlation matrix of ${}^pR_{\varepsilon\varepsilon}(0)$; ie., ${}^p\hat{R}_{\varepsilon\varepsilon}(0)$. An LU decomposition of this matrix would provide $L_{\gamma_p}$. The inverse matrix $L_{\gamma_p}^{-1}$ would then be the required estimate $\hat{L}_u^{-1}$.

Using eq(3.1b) in (3.8), we have

$$\underline{\gamma}_p(n) = L_{\gamma_p}^{-1}\underline{\varepsilon}_p(n) = L_{\gamma_p}^{-1}[\ \underline{x}(n) + \sum_{k=1}^{p} A_p^H(k)\ \underline{x}(n-k)] \tag{3.15a}$$

$$= \sum_{k=0}^{p} L_{\gamma_p}^{-1} A_p^H(k)\ \underline{x}(n - k). \tag{3.15b}$$

The block diagram for the filter implementation of eq(3.15a) is shown in Fig 3.2 as a two stage process with the first stage consisting of the conventional prediction error filter used to obtain $\underline{\varepsilon}_p(n)$ while the second stage consists of the operation on $\underline{\varepsilon}_p(n)$ with $L_{\gamma_p}^{-1}$ to yield the temporally and spatially uncorrelated process $\underline{\gamma}_p(n)$. The equivalent filter realization of eq(3.15b) is shown in Fig 3.2. From this viewpoint, the MMSE multichannel filter output which is both temporally whitened and 'spatially' whitened consists of coefficients $L_{\gamma_p}^{-1} A_p^H(k)$ k=0,1,...,p. In the next section, it is shown that these coefficients are contained in the matrix elements of $L_{\underline{\beta}}^{-1}$ where $L_{\underline{\beta}}$ is used in the $L_{\underline{\beta}}D_{\underline{\beta}}L_{\underline{\beta}}^H$ decomposition of $[R_{\underline{xx}}]$. Thus, the motivation for considering the $LDL^H$ decomposition is based on the anticipation of utilizing a single stage recursive procedure to obtain the filter coefficients $L_{\gamma_p}^{-1} A_p^H(k)$ required to estimate $\underline{\gamma}_p(n)$. In [1,14], the process $\underline{\gamma}_p(n)$ is

used in the development of a multichannel likelihood ratio for the detection of partially correlated random signals in noise.

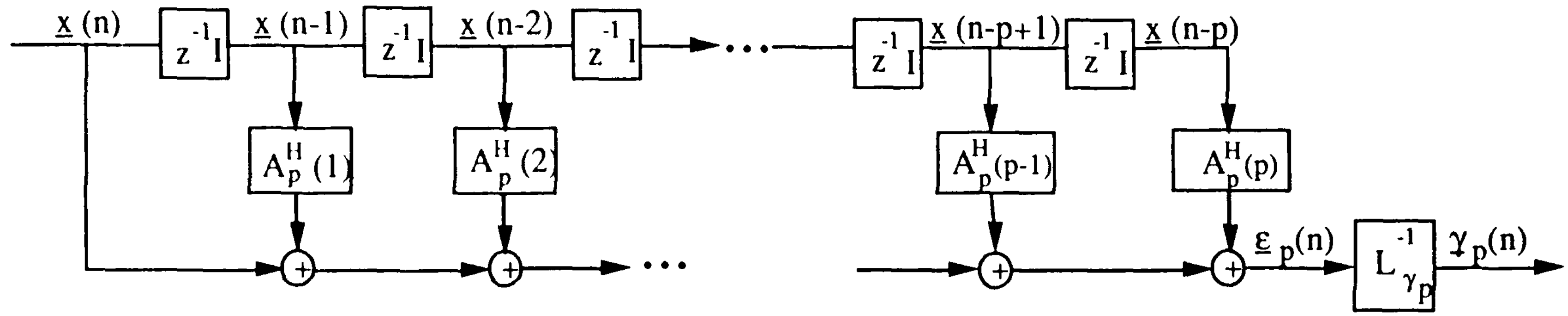


Figure 3.2 MULTI-CHANNEL PREDICTION ERROR FILTER (PEF)

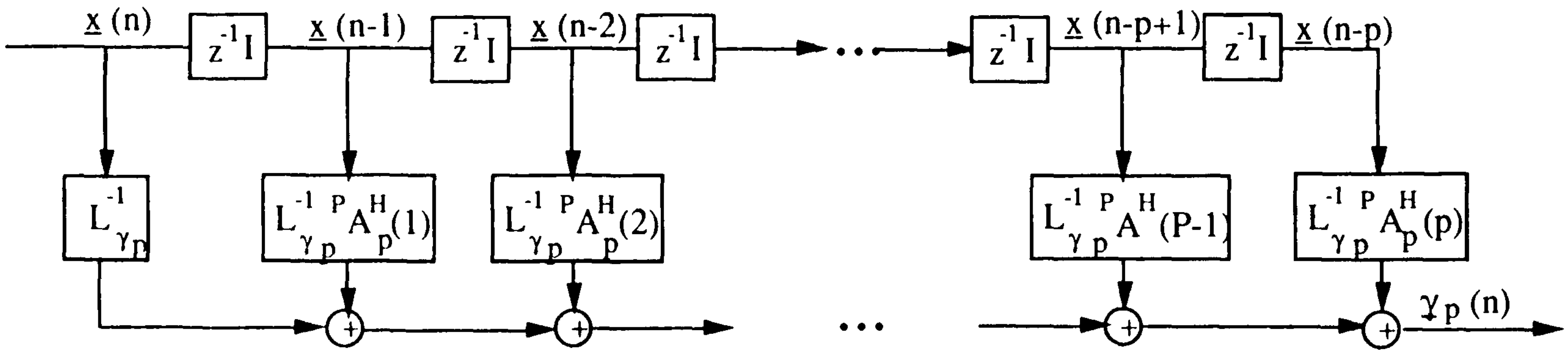


Figure 3.3 EQUIVALENT MULTI-CHANNEL PEF STRUCTURE

IV. Relation of Linear Prediction to Matrix Decomposition

In this section, it is shown that the coefficients of linear prediction for a multichannel random process are related to the multichannel correlation matrix through triangular decomposition. The discussion presented here represents a generalization of the scalar case noted in [3].

The JNxJN covariance matrix $R_{\underline{x}\underline{x}}$ for a set of N zero-mean, stationary random Jx1 vector processes $\underline{x}(n)$ n=1,2,...,N can be expressed using the form of eqs(2.15b) through (2.15d) with $\underline{\underline{x}}_{1,N}$ replacing eq(2.15c) and J indicating the number of channels. Since $R_{\underline{x}\underline{x}}$ is Hermitian and positive semi-definite, we can express it in terms of a unique factorization such that

$$R_{\underline{x}\underline{x}} = L_{\underline{\beta}} D_{\underline{\beta}} L_{\underline{\beta}}^H \tag{4.1}$$

where $L_{\underline{\beta}}$ is a JNxJN unit diagonal lower triangular matrix and $D_{\underline{\beta}}$ is a real, strictly diagonal JNxJN matrix (ie., with 1x1 diagonal elements). Solving eq(4.1) for $D_{\underline{\beta}}$, we obtain

$$D_{\underline{\beta}} = L_{\underline{\beta}}^{-1} R_{\underline{x}\underline{x}} (L_{\underline{\beta}}^{-1})^H \tag{4.2a}$$

$$= E\,[L_{\underline{\beta}}^{-1}\ \underline{\underline{x}}_{1,N}\underline{\underline{x}}_{1,N}^H (L_{\underline{\beta}}^{-1})^H] \tag{4.2b}$$

$$= E\,[\ \underline{\underline{\beta}}_{1,N}\underline{\underline{\beta}}_{1,N}^H] \tag{4.2c}$$

where

$$\underline{\underline{\beta}}_{1,N} = L_{\underline{\beta}}^{-1}\ \underline{\underline{x}}_{1,N} \tag{4.3}$$

contains uncorrelated elements in time and across channels. Writing eq(4.3) in expanded form

$$\underline{\underline{\beta}}_{1,N} = \begin{bmatrix} \underline{\beta}(1) \\ \underline{\beta}(2) \\ \underline{\beta}(3) \\ \cdots \\ \underline{\beta}(N) \end{bmatrix} = \begin{bmatrix} B_{11} & & & & \\ B_{21} & B_{22} & & [0] & \\ B_{31} & B_{32} & B_{33} & & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ B_{N1} & B_{N2} & B_{N3} & \cdots & B_{NN} \end{bmatrix} \begin{bmatrix} \underline{x}(1) \\ \underline{x}(2) \\ \underline{x}(3) \\ \\ \underline{x}(N) \end{bmatrix} \tag{4.4a}$$

where

$$\underline{\beta}^T(k) = [\beta_1(k)\ \beta_2(k)\ \dots\ \beta_J(k)] \qquad k = 1,2,\dots,N \tag{4.4b}$$

and $B_{m,n}$ is a JxJ matrix with $1 \leq m,n \leq N$ and $m \geq n$. $L_{\underline{\beta}}^{-1}$ is represented by the matrix in eq(4.4a). The matrices $B_{m,n}$ are full, in general, when m>n, but are lower triangular, unit diagonal when m=n [ie., when $B_{m,n}$ lies along the block diagonal of the matrix in eq(4.4a)]. It will be shown that each Jx(p+1)J block row of the matrix $L_{\beta}^{-1}$ contains the prediction error filter coefficients for a multichannel linear predictor of order p (for orders p=0 through N-1, respectively) which decorrelate the process in time and across channels.

It is now shown that these coefficients are equivalent to those obtained in the two stage whitening process developed in the previous section resulting in the temporally and spatially decorrelated error vector process, $\underline{\gamma}(n)$. This will be achieved by considering a two stage $LDL^H$ decomposition of $R_{\underline{XX}}$. For stationary processes, we can write eq(4.1) in block form as

$$R_{\underline{XX}} = R_{\underline{XX}}^B = \begin{bmatrix} R_{XX}(0) & R_{XX}(-1) & \dots & R_{XX}(-N+1) \\ R_{XX}(1) & R_{XX}(0) & \dots & R_{XX}(-N+2) \\ \dots & \dots & \dots & \dots \\ R_{XX}(N-1) & R_{XX}(N-2) & \dots & R_{XX}(0) \end{bmatrix} \tag{4.5}$$

where the JxJ correlation matrices on the RHS of eq(4.5) are expressed as

$$R_{XX}(l) = E[\underline{x}(n)\underline{x}^H(n - l)] \tag{4.6}$$

and the superscript B denotes block form. We note that although each $R_{XX}(l)$ in eq (4.5) is not Hermitian, $R_{\underline{XX}}^B$ is block Hermitian and positive semi-definite so that a block $LDL^H$ decomposition of eq(4.5) can be made. We therefore consider

$$R_{\underline{XX}}^B = L_{\underline{\varepsilon}}\, D_{\underline{\varepsilon}}\, L_{\underline{\varepsilon}}^H \tag{4.7}$$

where $L_{\underline{\varepsilon}}$ is a lower triangular matrix with JxJ identity matrices along the block diagonal and $D_{\underline{\varepsilon}}$ is block diagonal. Solving for $D_{\underline{\varepsilon}}$,

$$D_{\underline{\varepsilon}} = L_{\underline{\varepsilon}}^{-1} R_{\underline{XX}}^{B} (L_{\underline{\varepsilon}}^{-1})^{H} \tag{4.8a}$$

$$= L_{\underline{\varepsilon}}^{-1} E [ \underline{\underline{x}}_{1,N} \underline{\underline{x}}_{1,N}^{H} ] (L_{\underline{\varepsilon}}^{-1})^{H} \tag{4.8b}$$

$$= E [ L_{\underline{\varepsilon}}^{-1} \underline{\underline{x}}_{1,N} \underline{\underline{x}}_{1,N}^{H} (L_{\underline{\varepsilon}}^{-1})^{H} ] \tag{4.8c}$$

$$= E [ \underline{\underline{\varepsilon}}_{1,N} \underline{\underline{\varepsilon}}_{1,N}^{H} ] \tag{4.8d}$$

where

$$\underline{\underline{\varepsilon}}_{1,N} = L_{\underline{\varepsilon}}^{-1} \underline{\underline{x}}_{1,N}. \tag{4.9}$$

In Appendix B, the block rows of $L_{\underline{\varepsilon}}^{-1}$ are shown to contain the multichannel coefficients for forward linear prediction for filters of order p=N-1; ie.,

$$L_{\underline{\varepsilon}}^{-1} = \begin{bmatrix} I & & & & & \\ A_1^H(1) & I & & & [0] & \\ A_2^H(2) & A_2^H(1) & I & & & \\ A_3^H(3) & A_3^H(2) & A_3^H(1) & I & & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ A_p^H(p) & A_p^H(p-1) & A_p^H(p-2) & \cdots & A_p^H(1) & I \end{bmatrix} \tag{4.10}$$

From eq(4.8d)

$$D_{\underline{\varepsilon}} = E \left\{ \begin{bmatrix} \underline{\varepsilon}(1) \\ \underline{\varepsilon}(2) \\ \underline{\varepsilon}(3) \\ \cdots \\ \underline{\varepsilon}(N) \end{bmatrix} [ \underline{\varepsilon}^H(1) \, \underline{\varepsilon}^H(2) \ldots \underline{\varepsilon}^H(N) ] \right\} = \begin{bmatrix} {}^{0}R_{\varepsilon\varepsilon}(0) & & & [0] \\ & {}^{1}R_{\varepsilon\varepsilon}(0) & & \\ [0] & & \cdots & \\ & & & {}^{N-1}R_{\varepsilon\varepsilon}(0) \end{bmatrix} \tag{4.11}$$

where

$$^{p}R_{\varepsilon\varepsilon}(0) = E\,[\underline{\varepsilon}_p(n)\underline{\varepsilon}_p^H(n)] \equiv [\Sigma_f]_{\varepsilon_p} \qquad n = 1,2,\ldots,N \qquad p = N-1 \tag{4.12}$$

is the JxJ Hermitian covariance matrix of the error vector for the forward prediction error filter of order p.

We now perform an $LDL^H$ decomposition of $D_{\underline{\varepsilon}}$ as,

$$D_{\underline{\varepsilon}} = L_{\underline{\gamma}} D_{\underline{\Gamma}} L_{\underline{\gamma}}^H \tag{4.13}$$

where $L_{\underline{\gamma}}$ is a JNxJN unit diagonal lower triangular matrix and $D_{\underline{\Gamma}}$ is strictly diagonal (ie., with 1x1 diagonal elements). Solving for $D_{\underline{\Gamma}}$,

$$D_{\underline{\Gamma}} = L_{\underline{\gamma}}^{-1} D_{\underline{\varepsilon}} (L_{\underline{\gamma}}^{-1})^H. \tag{4.14}$$

From eq(4.8d)

$$D_{\underline{\Gamma}} = L_{\underline{\gamma}}^{-1} E\,[\underline{\varepsilon}_{1,N}\underline{\varepsilon}_{1,N}^H]\,(L_{\underline{\gamma}}^{-1})^H \tag{4.15a}$$

$$= E\,[L_{\underline{\gamma}}^{-1}\,\underline{\varepsilon}_{1,N}\underline{\varepsilon}_{1,N}^H\,(L_{\underline{\gamma}}^{-1})^H] \tag{4.15b}$$

$$= E\,[\,\underline{\gamma}_{1,N}\underline{\gamma}_{1,N}^H] \tag{4.15c}$$

where

$$\underline{\gamma}_{1,N} = L_{\underline{\gamma}}^{-1}\,\underline{\varepsilon}_{1,N} = L_{\underline{\gamma}}^{-1}\,L_{\underline{\varepsilon}}^{-1}\,\underline{x}_{1,N} \tag{4.16}$$

and the RHS of eq(4.16) follows from (4.9). We will now expand the matrix form of the RHS of eq(4.14). First, consider the form of $L_{\underline{\gamma}}^{-1}$ [the form of $L_{\underline{\varepsilon}}^{-1}$ is derived in Appendix B and noted in eq(4.10)]. Using the JxJ matrices $G_{m,n}$ $m \geq n$ for the off-diagonal block matrices of $L_{\underline{\gamma}}^{-1}$,

$$D_{\underline{\Gamma}} = \begin{bmatrix} G_{1,1} & [0] & [0] & \cdots & [0] \\ G_{2,1} & G_{2,2} & [0] & & \cdots \\ \cdots & & & \cdots & [0] \\ G_{N,1} & G_{N,2} & \cdots & G_{N,N-1} & G_{N,N} \end{bmatrix} \begin{bmatrix} {}^{0}R_{\varepsilon\varepsilon}(0) & & & [0] \\ & {}^{1}R_{\varepsilon\varepsilon}(0) & & \\ [0] & & \cdots & \\ & & & {}^{p}R_{\varepsilon\varepsilon}(0) \end{bmatrix}$$

$$\bullet \begin{bmatrix} G^H_{1,1} & G^H_{2,1} & \cdots & G^H_{N,1} \\ [0] & G^H_{2,2} & \cdots & G^H_{N,2} \\ \cdots & & \cdots & \cdots \\ [0] & \cdots & [0] & G^H_{N,N} \end{bmatrix} \qquad (4.17)$$

where p=N-1. In order to yield a strictly diagonal matrix $D_{\underline{\Gamma}}$, we must have

$$G_{m,n} = [0] \qquad \text{for all } m \neq n \qquad (4.18a)$$

and

$$[G_{p+1,p+1}]\ {}^{p}R_{\varepsilon\varepsilon}(0)\ [G_{p+1,p+1}]^H = D_{\gamma_p} \qquad (4.18b)$$

where $D_{\gamma_p}$ is the diagonal matrix for the pth block row of $D_{\underline{\Gamma}}$. Since $[G_{p+1,p+1}]$ is lower triangular, unit diagonal, and since eq(4.18b) implies $[G_{p+1,p+1}]\underline{\varepsilon}_p(n)$ is spatially uncorrelated, then

$$[G_{p+1,p+1}] = L^{-1}_{\gamma_p} \qquad (4.19)$$

where $L^{-1}_{\gamma_p}$ is lower triangular, unit diagonal. Eqs(4.18a) and (4.18b) imply that the first matrix in eq(4.17) can be written as

$$L_{\underline{\gamma}}^{-1} = \begin{bmatrix} L_{\gamma_0}^{-1} & & & [0] \\ & L_{\gamma_1}^{-1} & & \\ & & \cdots & \\ [0] & & & L_{\gamma_p}^{-1} \end{bmatrix}. \tag{4.20}$$

Therefore,

$$D_{\underline{\Gamma}} = \begin{bmatrix} L_{\gamma_0}^{-1}\,{}^{0}R_{\varepsilon\varepsilon}(0)L_{\gamma_0}^{-H} & & & [0] \\ & L_{\gamma_1}^{-1}\,{}^{1}R_{\varepsilon\varepsilon}(0)L_{\gamma_1}^{-H} & & \\ & & \cdots & \\ [0] & & & L_{\gamma_p}^{-1}\,{}^{p}R_{\varepsilon\varepsilon}(0)L_{\gamma_p}^{-H} \end{bmatrix} \tag{4.21a}$$

$$= \begin{bmatrix} D_{\gamma_0} & & & [0] \\ & D_{\gamma_1} & & \\ & & \cdots & \\ [0] & & & D_{\gamma_p} \end{bmatrix} \tag{4.21b}$$

and from eq(4.16)

$$\underline{\gamma}_{1,N} = \begin{bmatrix} L_{\gamma_0}^{-1} & & & [0] \\ & L_{\gamma_1}^{-1} & & \\ & & \cdots & \\ [0] & & & L_{\gamma_P}^{-1} \end{bmatrix} \begin{bmatrix} I & & & & & [0] \\ A_1^H(1) & I & & & & \\ A_2^H(2) & A_2^H(1) & I & & & \\ A_3^H(3) & A_3^H(2) & A_3^H(1) & I & & \\ A_P^H(P) & A_P^H(P-1) & A_P^H(P-2) & \cdots & A_P^H(1) & I \end{bmatrix} \begin{bmatrix} \underline{x}(1) \\ \underline{x}(2) \\ \underline{x}(3) \\ \underline{x}(4) \\ \cdots \\ \underline{x}(N) \end{bmatrix} \tag{4.22a}$$

so that

$$
\underline{y}_{1,N} = \begin{bmatrix} L_{\gamma_0}^{-1} & & & & \\ L_{\gamma_1}^{-1} A_1^H(1) & L_{\gamma_1}^{-1} & & [0] & \\ L_{\gamma_2}^{-1} A_2^H(2) & L_{\gamma_2}^{-1} A_2^H(1) & L_{\gamma_2}^{-1} & & \\ \cdots & \cdots & \cdots & \cdots & \\ L_{\gamma_P}^{-1} A_P^H(P) & L_{\gamma_P}^{-1} A_P^H(P-1) & L_{\gamma_P}^{-1} A_P^H(P-2) & \cdots \; L_{\gamma_P}^{-1} A_P^H(1) & L_{\gamma_P}^{-1} \end{bmatrix} \begin{bmatrix} \underline{x}(1) \\ \underline{x}(2) \\ \underline{x}(3) \\ \cdots \\ \underline{x}(N) \end{bmatrix} \tag{4.22b}
$$

$$
= L_{\underline{\Gamma}}^{-1} \, \underline{x}_{1,N}. \tag{4.22c}
$$

where $L_{\underline{\Gamma}}^{-1}$ is the matrix described in eq(4.22b). From the equality of eqs(4.16) and (4.22c), it follows that

$$
L_{\underline{\Gamma}}^{-1} = L_{\underline{\gamma}}^{-1} \, L_{\underline{\varepsilon}}^{-1}. \tag{4.23}
$$

Since $L_{\underline{\gamma}}^{-1}$ is unit diagonal lower triangular, each JxJ matrix $L_{\gamma_p}^{-1}$ is also unit diagonal lower triangular [see eq(4.20)]. And so, $L_{\underline{\Gamma}}^{-1}$ is unit diagonal lower triangular. Substituting eq(4.8a) into (4.14), we obtain

$$
D_{\underline{\Gamma}} = L_{\underline{\gamma}}^{-1} \, L_{\underline{\varepsilon}}^{-1} \, R_{\underline{XX}}^{B} \, (L_{\underline{\varepsilon}}^{-1})^H (L_{\underline{\gamma}}^{-1})^H \tag{4.24a}
$$

$$
= L_{\underline{\gamma}}^{-1} \, L_{\underline{\varepsilon}}^{-1} \, R_{\underline{XX}} \, (L_{\underline{\varepsilon}}^{-1})^H (L_{\underline{\gamma}}^{-1})^H \tag{4.24b}
$$

Using eq(4.23) in (4.24b),

$$
D_{\underline{\Gamma}} = L_{\underline{\Gamma}}^{-1} \, R_{\underline{XX}} \, (L_{\underline{\Gamma}}^{-1})^H \tag{4.25}
$$

From the uniqueness of the $LDL^H$ decomposition, it follows from eqs(4.2a) and (4.24) that

$$L_{\underline{\Gamma}}^{-1} = L_{\underline{\beta}}^{-1} \tag{4.26a}$$

and

$$D_{\underline{\Gamma}} = D_{\underline{\beta}}. \tag{4.26b}$$

And so, $L_{\underline{\Gamma}}^{-1}$ contains the multichannel coefficients of linear prediction for filters of order p=0 to N-1. Furthermore, these coefficients can be estimated either from an $LDL^H$ decomposition of $R_{\underline{xx}}$ or in a two stage operation as denoted by eq(4.16).

## V. Summary

In this report, we discussed the relationship of multichannel linear prediction to triangular matrix decomposition. Initially, a two stage prediction error filter is described which whitens the observation processes both in time and across channels. This filtering operation is such that it retains a causal form associated with linear prediction in both dimensions. For a filter of order p, these coefficients are $L_{\gamma_p}^{-1} A_p^H(k)$ k=0,1,...,p where $A_p^H(0)=I$, the JxJ identity matrix. Next, these coefficients are shown to be equivalent to those obtained through an $LDL^H$ decomposition of the multichannel correlation matrix. Specifically, the matrix coefficients contained in the matrix of eq(4.22b) are identical to those in eq(4.4a). The block rows of $L_{\Gamma}^{-1}$ and $L_{\beta}^{-1}$ contain the multichannel coefficients for linear prediction filters of orders zero through p=N-1 required to 'whiten' the observation process both in time and space (ie., across channels). It was shown that the prediction error process that is both temporally and spatially white can be obtained in one stage using the blocks from $L_{\beta}^{-1}$. The coefficient matrices $L_{\gamma_p}^{-1} A_p^H(k)$, k=0,1,...,p of the one-stage filter can also be obtained recursively by suitably modifying standard fast algorithms such as Levinson-Wiggins-Robinson, Nuttall-Strand and Morf-Vieira algorithms.

It is anticipated that recognition of the result derived in this report may lead to alternative processing methods involving recursive algorithms for multichannel signal processing applications as well as for two dimensional image processing [9].

References


[1] Michels, J. H., "A Parametric Detection Approach Using Multichannel Processes", RADC-TR-89-306, Nov. 1989, DTIC no. ADA - 219 -328.

[2] Michels, J. H., "Synthesis of multichannel autoregressive random processes and ergodicity considerations", RADC-TR-90-211, July, 1990.

[3] Therrien, C. W., "On the relation between triangular matrix decomposition and linear prediction", Proc. IEEE, vol 71, no.12, Dec. 1983, pp.1459-1460.

[4] Wiggins, R., Robinson, E., "Recursive solution to the multichannel filtering problem," JGR, Vol 70, 1965.

[5] Morf, M., Vierra, A. Lee, D., Kailath, T., "Recursive Multichannel maximum entropy spectral estimation", IEEE Trans. on Geoscience Electronic, vol GE-16, no 2, April 1978.

[6] Strand, O. N., "Multichannel complex maximum entropy (auto-regressive) spectral analysis", IEEE Trans. Antom. Control, vol, AC-22, pg 634-640, Aug 1977.

[7] Nuttall, A.H., FORTRAN Program for multivariate linear predictive spectral analysis, employing forward and backward averaging", Naval Underwater Systems Center TR-5419, New London, Conn., May 1976.

[8] Nuttall, A.H., "Multivariate linear predictive spectral analysis employing weighted forward and backward averaging: A generalization of Burg's algorithm, Naval Underwater Systems Center TR-5501, New London, Conn. Oct. 1976.

[9] Arun, K.S., Michels, J.H., "A tutorial on connections between Schur, Levinson, and Szego recursions for the 1-D multi-channel and 2-D single-channel linear prediction problems", invited presentation at the 24th Asilomar Conference on Signals, Systems, and Computers, 5-7, Nov. 1990.

[10] Therrien, C.W., "The Analysis of Multichannel Two-Dimensional Random Signals", Naval Postgraduate School technical report NPS62-87-002, 31 Oct. 1986.

[11] Haykin, S., Metford, P.A.S., "Experimental analysis of an innovations-based detection algorithm for surveillance radar," IEE Proc., Vol 132, Pt F, no 1, Feb. 1985.

[12] Metford, P.A.S., Haykin, S., Taylor, D., "An innovations approach to discrete-time detection theory," IEEE Trans. on Information Theory, Vol IT-28, no 2, March 1982.

[13] Haykin, S., Metford, P.A.S., "Some limiting forms of an Innovations-Based Discrete-Time Detection Algorithm," IEEE Trans. on Aerospace and Electronic Systems, AES-23, no.3, May 1987.

[14] Michels, J.H., "On Model-Based Multichannel Detection," Ph. D. thesis in progress.

APPENDIX A

In this section, we determine the conditions under which $\underline{\varepsilon}_p(n)$ as expressed in eq(3-1) is a temporal white noise MMSE output of a linear prediction error filter. The linear prediction error of x(n) as defined in eq (3.1a) is

$$\underline{\varepsilon}_p(n) = \underline{x}(n) - \hat{\underline{x}}(n|n-1) \tag{A.1}$$

where $\hat{x}(n|n-1)$ represents the estimated vector of x(n) using past data values with the initial condition x(1|0) = 0. Using a linear predictor with p past values, we define

$$\hat{\underline{x}}(n|n-1) = -\sum_{k=1}^{p} A_p^H(k)\underline{x}(n-k) \tag{A.2}$$

where $A_p^H(k)$, k = 1,2,...,p are JxJ matrices representing the coefficients of the linear predictor. Substituting eq (A-2) in (A-1)

$$\underline{\varepsilon}_p(n) = \underline{x}(n) + \sum_{k=1}^{p} A_p^H(k)\underline{x}(n-k) \tag{A-3a}$$

$$= \sum_{k=0}^{p} A_p^H(k)\underline{x}(n-k) \tag{A-3b}$$

where $A_p^H(0) = I$. Let the concatenated column vector of p+1 vectors (each of dimension J) be defined as

$$\underline{\underline{x}}^T_{n,n-p} = [\underline{x}^T(n)\ \underline{x}^T(n-1)...\underline{x}^T(n-p). \tag{A.4}$$

Post multiplying eq (A-3) by $\underline{\underline{x}}^H_{n,n-p}$ and taking the expected value

$$E[\underline{\varepsilon}(n)\underline{\underline{x}}^H_{n,n-p}] = E[A_p^H\underline{\underline{x}}_{n,n-p}\underline{\underline{x}}^H_{n,n-p}] \tag{A.5a}$$

$$= A_p^H E[\underline{\underline{x}}_{n,n-p}\underline{\underline{x}}^H_{n,n-p}] \tag{A.5b}$$

$$= A_p^H[\tilde{R}_{xx}] \tag{A.5c}$$

where $A_p^H$ and the reversed order correlation matrix $[\tilde{R}_{xx}]$ are defined in eq(2.12a) and the discussion following (2.12b), respectively. Eq (A.5) can also be written such that

$$E[\underline{\varepsilon}_p(n)\underline{\underline{x}}_{n,n-p}^H] = E\left\{\underline{\varepsilon}_p(n)\left[\underline{x}^H(n)\,\underline{x}^H(n-1)\,\underline{x}^H(n-2)\,...\,\underline{x}^H(n-p)\right]\right\} \tag{A.6}$$

We now determine the coefficients of $A_p^H$ subject to the condition that $\underline{\varepsilon}(n)$ is a MMSE vector. Under this condition,

$$E\,[\underline{\varepsilon}_p(n)\,\underline{x}^H(n-k)] = [0] \qquad k>0. \tag{A.7}$$

Eq (A.7) is the orthogonality condition which states that the error vector is orthogonal to past observation values. Using this condition, eq (A.6) becomes

$$E[\underline{\varepsilon}_p(n)\underline{\underline{x}}_{n,n-p}^H] = \left\{E\,[\underline{\varepsilon}_p(n)\underline{x}^H(n)]\,[0]\,[0]\,...\,[0]\right\} \tag{A.8}$$

From eq (A.3a)

$$\underline{x}(n) = \underline{\varepsilon}_p(n) - \sum_{k=1}^{p} A_p^H(k)\underline{x}(n-k) \tag{A.9}$$

so that taking the Hermitian transpose

$$\underline{x}^H(n) = \underline{\varepsilon}_p^H(n) - \sum_{k=1}^{p} \underline{x}^H(n-k)A_p(k). \tag{A.10}$$

Using eq (A.10) in the RHS of eq (A.8)

$$E[\underline{\varepsilon}_p(n)\underline{\underline{x}}_{n,n-p}^H] = \left\{\left[E[\underline{\varepsilon}_p(n)\underline{\varepsilon}_p^H(n)] - \sum_{k=1}^{p} E[\underline{\varepsilon}_p(n)\underline{x}^H(n-k)]A_p(k)\right][0]...[0]\right\} \tag{A.11a}$$

$$=\{[\Sigma_f]_{\varepsilon_p}\;[0]\,[0]...[0]\} \tag{A.11b}$$

where we have again used eq (A.7) to yield eq (A.11) and $[\Sigma_f]_{\varepsilon_p}$ is the forward prediction error covariance matrix for the filter of order p. Combining eq (A-5) and (A-11), we have

$$A_p^H[\tilde{R}_{xx}] = \{[\Sigma_f]_{\varepsilon_p}\ [0]\ [0]...[0]\}. \tag{A-12}$$

Eq (A.12) is the multichannel AR Yule-Walker normal equation in augmented form. It provides a set of JP linear equations to solve for the values of the matrix coefficients which minimize the mean square error vector. Although eq (A-12) has often been presented in the literature, the reversed order form of the correlation matrix has not often been noted [10]. We will utilize this feature in Appendix B.

We now show that the vector process $\underline{\varepsilon}_p(n)$ is uncorrelated in time. At an arbitrary time (n-l) where l>0, eq (A-10) becomes

$$\underline{x}^H(n-l) = \underline{\varepsilon}_p^H(n-l) - \sum_{k=1}^{p} \underline{x}^H(n-k-l)A_p(k). \qquad l > 0. \tag{A.13}$$

Using eq (A.13) in eq (A.7)

$$E[\underline{\varepsilon}_p(n)\underline{x}^H(n-l)] = E[\underline{\varepsilon}_p(n)\underline{\varepsilon}_p^H(n-l)] - \sum_{k=1}^{p} E[\underline{\varepsilon}_p(n)\underline{x}^H(n-k-l)]A_p(k) = [0]. \tag{A.14}$$

From eq (A7), we have

$$E\,[\underline{\varepsilon}_p(n)\underline{x}^H(n-k-l) = [0] \qquad k>0. \tag{A.15}$$

And so, the summation term in eq(A.14) is zero so that

$$E\,[\underline{\varepsilon}_p(n)\underline{\varepsilon}_p^H(n-l)] = [0] \qquad l>0. \tag{A.16}$$

Thus, the sequence of vector outputs from the MMSE prediction error filter are orthogonal. Since $\underline{\varepsilon}_p(n)$ is a zero mean Gaussian process, its sequence of values $\{\varepsilon_p(n)\}$ are mutually independent so that $\{\underline{\varepsilon}_p(n)\}$ is a white, Gaussian noise sequence. However, these processes are, in general, correlated across the channels.

APPENDIX B

In this appendix, we show that the matrix coefficients of multichannel linear prediction for a multichannel random process and the prediction error covariance matrices are related to the covariance matrix through a block triangular decomposition. The procedure is a straight forward generalization of [3]. In this discussion, we are considering the first stage of processing as noted in sections III and IV. This stage results in the $\underline{\varepsilon}_p(n)$ vector output. Therefore, we only consider the $A_p(k)$ matrix coefficients here. A treatment which includes the second stage of processing using the matrix $L_{\gamma_p}$ to obtain $\underline{\gamma}_p(n)$ is developed in Section IV.

Recognizing that $R^B_{\underline{xx}}$ defined in eq(4.5) is Hermitian positive semi-definite, we can obtain

$$R^B_{\underline{xx}} = L_{\underline{\varepsilon}} D_{\underline{\varepsilon}} (L_{\underline{\varepsilon}})^H \tag{B.1}$$

where $L_{\underline{\varepsilon}}$ is lower block triangular with the identity matrix I forming the block diagonal matrices and $D_{\underline{\varepsilon}}$ is a real block diagonal matrix. Solving eq(B.1) for $D_{\underline{\varepsilon}}$

$$D_{\underline{\varepsilon}} = L^{-1}_{\underline{\varepsilon}} R^B_{\underline{xx}} (L^{-1}_{\underline{\varepsilon}})^H \tag{B.2}$$

If we consider

$$\underline{\underline{\varepsilon}}_{1,N} = L^{-1}_{\underline{\varepsilon}} \underline{\underline{x}}_{1,N} \tag{B.3}$$

we can easily show that

$$D_{\underline{\varepsilon}} = E\,[\underline{\underline{\varepsilon}}_{1,N}\, \underline{\underline{\varepsilon}}^H_{1,N}\,] \tag{B.4}$$

where

$$\underline{\underline{\varepsilon}}^T_{1,N} = [\underline{\varepsilon}^T(1)\; \underline{\varepsilon}^T(2) \ldots \underline{\varepsilon}^T(N)] \tag{B.5a}$$

and

$$\underline{\underline{x}}^T_{1,N} = [\underline{x}^T(1)\; \underline{x}^T(2) \ldots \underline{x}^T(N)] \tag{B.5b}$$

Since $L_\varepsilon$ is lower block triangular with unit diagonal elements, $L_\varepsilon^{-1}$ has the same form so that eq (B.3) is a causal and causally invertible transformation of the data. We now consider the normal equations for a multichannel predictor of order p such that

$$A_p^H[\tilde{R}_{\underline{XX}}] = [\Sigma_f]_{\varepsilon_p}^H \underline{I}^T \qquad \text{(B-6)}$$

where $[\tilde{R}_{\underline{XX}}]_p$ is the reversed order multichannel covariance matrix,

$$A_p^H = [I\ A_p^H(1)\ A_p^H(2) \ldots A_p^H(p)] \qquad \text{(B.7)}$$

and

$$\underline{I}^T = \{I\ [0]\ [0] \ldots [0]\} \qquad \text{(B.8)}$$

where I is a JxJ identity matrix. The vector of matrices $\underline{A}_p^H$ is the vector of multichannel pth order linear prediction coefficients and $[\Sigma_f]_{\varepsilon_p}$ is the Hermitian, pth order, multichannel forward prediction error covariance matrix. Post multiplying eq (B.6) by $A_p$, and recognizing that $[\Sigma_f]_{\varepsilon_p}$ is Hermitian, we obtain

$$[I\ A_p^H(1)\ A_p^H(2) \ldots A_p^H(p)][\tilde{R}_{\underline{XX}}]\begin{bmatrix} I \\ A_p(1) \\ A_p(2) \\ \vdots \\ A_p(p) \end{bmatrix} = [\Sigma_f]_{\varepsilon_p} \qquad \text{(B.9)}$$

Using the relation

$$G_B^{-1}\, G_B = I \qquad \text{(B.10)}$$

where $G_B$ is the block reflection matrix (i.e., the square J(p+1) x J(p+1) matrix with JxJ identity matrices along the block cross diagonal), we have

$$[I\ A_p^H(1)\ A_p^H(2) \ldots A_p^H(p)]\, G_B^{-1} = [A_p^H(p) \ldots A_p^H(2)\ A_p^H(1)\ I\ ] \qquad \text{(B.11a)}$$

and

$$G_B\,[\tilde{R}_{\underline{xx}}]G_B^{-1} = [R_{\underline{xx}}]. \tag{B.11b}$$

Using these expressions in (B.9), we have

$$[A_p^H(p)\ldots A_p^H(2)\ A_p^H(1)\ I][R_{\underline{xx}}]\begin{bmatrix} A_p(p) \\ \vdots \\ A_p(2) \\ A_p(1) \\ I \end{bmatrix} = [\Sigma_f]_{\varepsilon_p} \tag{B.12}$$

We now write eq (B.12) for p = 0, 1, ..., N-1 in the combined form as

$$\begin{bmatrix} I & & & & & \\ A_1^H(1) & I & & & [0] & \\ A_2^H(2) & A_2^H(1) & I & & & \\ A_3^H(3) & A_3^H(2) & A_3^H(1) & I & & \\ \ldots & \ldots & \ldots & \ldots & \ldots & \\ A_p^H(p) & A_p^H(p-1) & A_p^H(p-2) & \ldots & A_p^H(1) & I \end{bmatrix}\begin{bmatrix} R_{xx}(0) & R_{xx}(-1) & \ldots & R_{xx}(-N+1) \\ R_{xx}(1) & R_{xx}(0) & \ldots & R_{xx}(-N+2) \\ R_{xx}(2) & R_{xx}(1) & R_{xx}(0) & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ R_{xx}(N-1) & R_{xx}(N-2) & \ldots & R_{xx}(0) \end{bmatrix}$$

$$\bullet\begin{bmatrix} I & A_1(1) & A_2(2) & A_3(3) & \ldots & A_p(p) \\ & I & A_2(1) & A_3(2) & \ldots & A_p(p-1) \\ & & I & A_3(1) & \ldots & A_p(p-2) \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ [0] & & & & & I \end{bmatrix} = \begin{bmatrix} [\Sigma_f]_{\varepsilon_0} & & & & \\ & [\Sigma_f]_{\varepsilon_1} & & [0] & \\ & & [\Sigma_f]_{\varepsilon_2} & & \\ & [0] & & \ldots & \\ & & & & [\Sigma_f]_{\varepsilon_{N-1}} \end{bmatrix} \tag{B.13}$$

Eq (B.13) is of the same form as eq (B.2). Since the causal decomposition in eq(B.2) is unique† , then $L_{\underline{\varepsilon}}^{-1}$ can be identified with the lower triangular matrix in eq (B.13). Thus the block rows of $L_{\underline{\varepsilon}}^{-1}$ are the multichannel coefficients for

† The uniqueness of this decomposition is based upon the specified block form $R_{\underline{xx}}^{B}$ defined using the concatenated vector in eq(B.5b). However, other block forms of $R_{\underline{xx}}$ could have been made which still retain the Hermitian property.

linear predictive filters of orders zero through p=N-1 and the block diagonal matrix elements of $D_{\varepsilon}$ are the prediction error variances associated with the multichannel filter orders.